PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : C07K 13/00, A61K 37/02 | A1

(11) International Publication Number: WO 89/ 05825

(43) International Publication Date: 29 June 1989 (29.06.89)

(21) International Application Number: PCT/US88/04659

(22) International Filing Date: 21 December 1988 (21.12.88)

(31) Priority Application Number: 137,557

(32) Priority Date: 22 December 1987 (22.12.87)

(33) Priority Country: US

(71) Applicant: IMMUNEX CORPORATION [US/US]; 51 University Street, Seattle, WA 98101 (US).

(72) Inventors: BOOTH, Roger, J. ; 9 Mozley Avenue, Devonport, Auckland (NZ). HARRIS, David, P. ; 2/6 Tarura Terrace, Parnell, Auckland (NZ). LOVE, Jennifer, M. ; 31 Northumberland Avenue, Takapuna, Auckland (NZ). WATSON, James D. ; 769 Riddel Road, St. Heliers (NZ).

(74) Agent: HALLQUIST, Scott, G.; Immunex Corporation, 51 University Street, Seattle, WA 98101 (US).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: EXPRESSION OF RECOMBINANT 18/16.6 AND 15.2 kD PROTEIN ANTIGENS OF *MYCOBACTERIUM LEPRAE*

(57) Abstract

DNA sequences encoding *Mycobacterium leprae*-specific antigens are disclosed. These polypeptides are useful in diagnosis of leprosy by immunoassay and in vaccines for prevention of the disease.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	ML	Mali
AU	Australia	GA	Gabon	MR	Mauritania
BB	Barbados	GB	United Kingdom	MW	Malawi
BE	Belgium	HU	Hungary	NL	Netherlands
BG	Bulgaria	IT	Italy	NO	Norway
BJ	Benin	JP	Japan	RO	Romania
BR	Brazil	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SN	Senegal
CH	Switzerland	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
DE	Germany, Federal Republic of	LU	Luxembourg	TG	Togo
DK	Denmark	MC	Monaco	US	United States of America
FI	Finland	MG	Madagascar		

TITLE

Expression of Recombinant 18/16.6 and 15.2 kD Protein Antigens of Mycobacterium leprae

BACKGROUND OF THE INVENTION

Leprosy is a chronic infectious disease caused by the acid-fast bacillus Mycobacterium leprae which exhibits a spectrum of clinical manifestations between extreme lepromatous and tuberculoid forms. The disease afflicts from 10 to 15 million people worldwide and is a major health problem in many tropical and sub-tropical areas.

Although leprosy can be treated with chemotherapy, successful control of the disease will ultimately depend upon effective vaccination. As M. leprae is an obligate intracellular parasite, vaccination procedures must be directed preferentially towards stimulation of cell-mediated responses. Initial results using a vaccine comprised of a mixture of armadillo-derived, killed M. leprae organisms and M. bovis BCG (bacillus Calmette-Guerin) have shown promise, and more extensive preventative trials are now underway using this combination. Unfortunately, it is questionable whether such vaccines, should they prove efficacious, would be inexpensive or plentiful enough to permit widespread use. The development of recombinant vaccines offers a potential solution to this problem.

Recently, Young et al. (Nature 316:450, 1985) constructed a random library of M. leprae DNA sequences in the λgt11 bacteriophage expression vector and, using a panel of monoclonal antibodies directed against antigenic proteins and glycoproteins of M. leprae, were able to isolate recombinant clones expressing antigenic epitopes of the five most immunogenic M. leprae protein antigens. The sequence of one such antigen, termed the 65 kilodalton (kDa) antigen, is disclosed in PCT application US 87/01825, filed July 28, 1987. Crude lysate material from those recombinants which expressed an epitope of an 18 kilodalton (kDa) M. leprae protein was found to stimulate proliferation in culture of some M. leprae-specific human T cell clones produced from M. leprae-vaccinated volunteers (Mustafa et al., Nature 319:63, 1986). Most of the antigenic proteins of M. leprae exhibit immunological cross-reactivity at the level of antibodies with proteins from other mycobacteria. Significantly, however, the 18 kDa

protein appears to include at least one epitope specific for M. leprae, as shown by binding to a monoclonal antibody designated L5 (Britton et al., J. Immunol. 135:4171, 1986).

The 18 kDa protein is likely, therefore, to contain peptide fragments of potential importance as vaccine candidates near, or coincident with, the epitope recognized by the monoclonal antibody. The present invention is based upon the isolation and sequencing of a full length gene encoding the M. leprae 18 kDa protein from an M. leprae genomic library. The predicted molecular weight of the 18 kDa protein, based upon its apparent amino acid sequence, is actually 16.6 kDa. During screening of the λgt11 library described by Young et al., supra, it was determined that this library lacked the complete sequence of the 18 kDa antigen due to an unexpected cleavage by a restriction endonuclease during preparation of the library. Accordingly, an additional source of M. leprae DNA was screened to secure the complete sequence, which was determined as set forth in the accompanying specification.

In the course of this work, a second open reading frame was discovered downstream from that encoding the 18/16.6 kDa antigen. This sequence apparently encodes a 15.2 kDa polypeptide representing a potential M. leprae antigen. Expression of these sequences in recombinant systems can provide substantial quantities of M. leprae-specific antigen at a reasonable cost for use in diagnosis or treatment. In addition, knowledge of the amino acid sequence of these polypeptides provides a basis for determining whether they contain any peptide regions likely to be good epitopes for helper T cell stimulation. As there is increasing evidence for suppressor T cell involvement in leprosy, it will be vital to discover whether particular domains of M. leprae-specific proteins, for example, the 18/16.6 kDa protein, can be identified that are capable of immunizing M. leprae-specific T helper cells as opposed to suppressor cells.

SUMMARY OF THE INVENTION

The present invention provides recombinant expression vectors comprising DNA sequences encoding polypeptides which are substantially identical to the 18/16.6 kDa and 15.2 kDa protein antigens of Mycobacterium leprae, and substantially homogeneous protein

compositions consisting essentially of these polypeptides. These polypeptides, and immunologically effective subunits thereof, are useful immunoassay reagents for diagnosis of leprosy, and are also potentially useful as immunogenic components of vaccines.

BACKGROUND OF THE INVENTION

Prior to the present invention, the DNA sequences of the 18/16.6 kDa and 15.2 kDa protein antigens disclosed herein were unknown. Although others had prepared DNA libraries comprising such sequences, and the 18/16.6 kDa protein had actually been expressed in lysates of E. coli infected with λgt11 bacteriophage, the expression products thus provided were actually fusion proteins having extensive E. coli β-galactosidase domains. For use as immunoassay reagents or for use in vaccines, substantial quantities of highly purified antigen are desirable. To ensure effectiveness in diagnosis or vaccination, expression of recombinant antigens without an extensive foreign protein portion is preferable.

Recombinant clones expressing antigenic determinants of the 18/16.6 kDa protein antigen from M. leprae which is recognized by the L5 monoclonal antibody (Mustafa et al., Nature 319:63, 1986) were isolated and sequenced. All clones expressed the M. leprae-specific determinant as part of a large fusion protein having an E. coli β-galactosidase fragment. The deduced amino acid sequence of the coding region indicated that the λgt11 recombinant clones contained an incomplete M. leprae gene sequence representing the carboxy-terminal two-thirds (111 amino acids) of the 18 kDa gene and coding for a peptide of molecular weight 12,432 daltons (Da). Subsequent isolation and sequencing of a 3.2 kb BamHI-PstI DNA fragment from a genomic M. leprae cosmid library permitted the the complete 148 amino acid sequence to be determined. This sequence encoded a polypeptide having a predicted molecular weight of 16,607 Da. Accordingly, it will hereinafter be referred to as the "18/16.6 kDa" antigen. The residues providing the epitope recognized by antibody L5 are found in the sequence of amino acids 101-116 depicted in Table 1, particularly residues 109-114.

A second open reading frame 560 bases downstream (3' with respect to the first open reading frame described above) was found to

code for a putative protein of 137 amino acids having a predicted molecular weight of 15,196 Da. Neither this protein nor the 18/16.6 kDa amino acid sequence displayed any significant homologies with any proteins listed in the GENBANK, EMBL or NBRF databases.

Definitions

"18/16.6 and 15.2 kDa protein antigens of Mycobacterium leprae" refer to proteins expressed by pathogenic strains of M. leprae having the amino acid sequences which are predicted from the nucleotide sequences depicted in Figures 1 and 2. "Immunologically effective subunit" means a peptide sequence of sufficient length to provide an epitope capable of recognition by antibody.

"Substantially identical", when used to define amino acid sequences, means that a particular subject sequence, for example, a mutant sequence, varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between reference and subject sequences. For purposes of the present invention, amino acid sequences having greater than 80 percent identity are considered to be substantially identical. In defining nucleic acid sequences, all subject nucleic acid sequences capable of encoding substantially identical amino acid sequences are considered substantially identical to a reference nucleic acid sequence. For purposes of determining substantial identity, truncations or internal deletions of the reference sequence should be disregarded. Sequences having lesser degrees of similarity, comparable biological activity, and equivalent expression characteristics are considered to be equivalents.

"Substantially homogeneous protein composition" refers to a preparation consisting essentially of a selected polypeptide antigen of M. leprae which is free of detectable quantities of other M. leprae proteins. Such compositions are conveniently provided as a product of recombinant microbial expression systems, for example, bacteria or yeast.

"DNA sequence" refers to a DNA polymer, in the form of a separate fragment or as a component of a larger DNA construct, which has been derived from DNA isolated at least once in substantially pure form, i.e., free of contaminating endogenous materials and in a

quantity or concentration enabling identification, manipulation, and recovery of the sequence and its component nucleotide sequences by standard biochemical methods, for example, using a cloning vector. Sequences of non-translated DNA may be present 5' or 3' from the open reading frame, where the same do not interfere with manipulation or expression of the coding regions.

"Nucleotide sequence" refers to a heteropolymer of deoxyribonucleotides. DNA sequences encoding the proteins provided by this invention can be assembled from cDNA fragments and short oligonucleotide linkers, or from a series of oligonucleotides, to provide a synthetic gene which is capable of being expressed in a recombinant transcriptional unit.

"Recombinant expression vector" refers to a plasmid comprising a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription and translation initiation and termination sequences. Structural elements intended for use in yeast expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an N-terminal methionine residue. This residue may optionally be subsequently cleaved from the expressed recombinant protein to provide a final product.

"Recombinant microbial expression system" means a substantially homogeneous monoculture of suitable host microorganisms, for example, bacteria such as E. coli or yeast such as S. cerevisiae, which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit as a component of a resident plasmid. Generally, cells constituting the system are the progeny of a single ancestral transformant. *Recombinant expression systems* as defined herein will express heterologous protein upon induction of the regulatory elements linked to the DNA sequence or synthetic gene to be expressed.

Appropriate cloning and expression vectors for use with bacterial and fungal hosts are described by Pouvels et al. (Cloning

Vectors: A Laboratory Manual, Elsevier, New York, 1985), the relevant disclosure of which is hereby incorporated by reference.

Useful expression vectors for bacterial use can be constructed by inserting a DNA sequence encoding a selected polypeptide together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure amplification within the host. Suitable prokaryotic hosts for transformation include E. coli, Bacillus subtilis, Salmonella typhimurium, and various species within the genera Pseudomonas, Streptomyces, and Staphylococcus, although others may also be employed as a matter of choice.

Expression vectors are conveniently constructed by cleavage of cDNA clones at sites close to the codon encoding the N-terminal residue of the mature protein. Synthetic oligonucleotides can then be used to "add back" any deleted sections of the coding region and provide a linking sequence for ligation of the coding fragment in appropriate reading frame in the expression vector, and optionally a codon specifying an initiator methionine.

Useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and pGEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

One useful bacterial expression system employs the phage λP_L promoter and cI857 thermolabile repressor. Plasmid vectors available from the American Type Culture Collection which incorporate derivatives of the λP_L promoter include plasmid pHUB2, resident in E. coli strain JMB9 (ATCC 37092) and pPLc28, resident in E. coli RR1 (ATCC 53082). Other useful promoters for expression in E. coli include the T7 RNA polymerase promoter described by Studier et al. (J. Mol. Biol. 189: 113, 1986), the lacZ promoter described by Lauer (J. Mol. Appl. Genet. 1:139-147, 1981) and available as ATCC 37121, and the tac promoter described by Maniatis (Molecular Cloning: A

Laboratory Manual, Cold Spring Harbor Laboratory, 1982, p 412) and available as ATCC 37138.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification. Cells are grown, for example, in a 10 liter fermenter employing conditions of maximum aeration and vigorous agitation. An antifoaming agent (Antifoam A) is preferably employed. Cultures are grown at 30°C in the superinduction medium disclosed by Mott et al. (Proc. Natl. Acad. Sci. USA 82:88, 1985), optionally including antibiotics, derepressed at a cell density corresponding to A_{600} = 0.4-0.5 by elevating the temperature to 42°C, and harvested from 2-20, preferably 3-6, hours after the upward temperature shift. The cell mass is initially concentrated by filtration or other means, then centrifuged at 10,000 x g for 10 minutes at 4°C followed by rapidly freezing the cell pellet.

Yeast systems, preferably employing Saccharomyces species such as S. cerevisiae, can also be employed for expression of the recombinant proteins of this invention. Yeast of other genera, for example, Pichia or Kluyveromyces, have also been employed as production strains for recombinant proteins.

Generally, useful yeast vectors will include origins of replication and selectable markers permitting transformation of both yeast and E. coli, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed yeast gene to induce transcription of a downstream structural sequence. Such promoters can be derived from yeast transcriptional units encoding highly expressed gene such as 3-phosphoglycerate kinase (PGK), α-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate reading frame with translation initiation and termination sequences, and, preferably, a leader sequence capable of directing secretion of translated protein into the extracellular medium. Optionally, the heterologous sequence can encode a fusion

protein including an N-terminal identification peptide (e.g., Asp-Tyr-Lys-$(Asp)_4$-Lys) or other sequence imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

Useful yeast vectors can be assembled using DNA sequences from pBR322 for selection and replication in E. coli (Amp^r gene and origin of replication) and yeast DNA sequences including a glucose-repressible alcohol dehydrogenase 2 (ADH2) promoter. The ADH2 promoter has been described by Russell et al. (J. Biol. Chem. 258:2674, 1982), and Beier et al. (Nature 300:724, 1982). Such vectors may also include a yeast TRP1 gene as a selectable marker and the yeast 2 μ origin of replication. A yeast leader sequence, for example, the α-factor leader which directs secretion of heterologous proteins from a yeast host, can be inserted between the promoter and the structural gene to be expressed (see Kurjan et al., U.S. Patent 4,546,082; Kurjan et al., Cell 30:933 (1982); and Bitter et al., Proc. Natl. Acad. Sci. USA 81:5330, 1984). The leader sequence may be modified to contain, near its 3' end, one or more useful restriction sites to facilitate fusion of the leader sequence to foreign genes.

Suitable yeast transformation protocols are known to those skilled in the art; an exemplary technique is described by Hinnen et al. (Proc. Natl. Acad. Sci. USA 75:1929, 1978), selecting for Trp^+ transformants in a selective medium consisting of 0.67% yeast nitrogen base, 0.5% casamino acids, 2% glucose, 10 μg/ml adenine and 20 μg/ml uracil.

Host strains transformed by vectors comprising the ADH2 promoter may be grown for expression in a rich medium consisting of 1% yeast extract, 2% peptone, and 1% glucose supplemented with 80 μg/ml adenine and 80 μg/ml uracil. Derepression of the ADH2 promoter occurs upon exhaustion of medium glucose. Crude yeast supernatants are harvested by filtration and held at 4°C prior to further purification.

Recombinant proteins can be purified from crude extracts or crude fermentates by various conventional chromatographic methods, optionally following concentration by ultrafiltration or precipitation with ammonium sulfate. Chromatography can be by size-exclusion, ion-exchange, or affinity chromatography using a specific antibody conjugated to a solid phase. Finally, one or more reversed-phase high

performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the recombinant protein.

In one embodiment of the present invention, the amino acid sequence of an M. leprae antigen is linked to a yeast α-factor leader sequence or bacterial signal sequence via an N-terminal fusion construct comprising a nucleotide encoding the peptide Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys (DYKDDDDK). The latter sequence is highly antigenic and provides an epitope reversibly bound by specific monoclonal antibody (4E11, ATCC HB-9259), enabling rapid assay and facile purification of expressed recombinant protein. This sequence is also specifically cleaved by bovine mucosal enterokinase at the residue immediately following the Asp-Lys pairing. Fusion proteins capped with this peptide may also be resistant to intracellular degradation in E. coli. The presence of the DYKDDDDK epitope in fusion proteins comprising other domains recognized by M. leprae-specific antibody also permits the fusion proteins to be used in two-site immunoassays.

Utility

The protein antigens provided by the present invention represent useful components of diagnostic immunoassays for antibodies specific to M. leprae. Methods of conducting such assays are well known to those of skill in the art. One useful method is an indirect two-site immunoassay, also known as an "antibody sandwich" assay. In this method, a purified antibody ("first antibody") which specifically recognizes an epitope present on the antigen to be employed is nonspecifically bound to a solid phase support, for example, a plastic microtiter dish or plastic tube. This preparation is then reacted with purified antigen to provide a complex of antigen and first antibody bound to the solid phase, and excess antigen is removed by decanting and washing. Various dilutions of samples (for example, human sera) to be tested for the presence of antibody specific for the antigen ("second antibody") are then reacted with the antigen-first antibody complex. Excess sample is removed, and the presence of second antibody bound to the complex is detected using reagents

specific for human IgG. Such reagents include, for example, rabbit or goat anti-human IgG polyclonal antibodies conjugated to a suitable enzymatic or radiometric reporter group, for example, alkaline phosphatase, horseradish peroxidase, or ^{125}I. The presence of the reporter group is then detected by appropriate means.

Various technologies applicable to the design of immunoassays are described by Hales et al., Methods Enzymol. 70:334 (1980) and the references cited therein. Reagents and supplies employed in such assays are widely commercially available.

The protein antigens and peptide subunits of the present invention are also suitable for use as vaccines. The vaccines can comprise the entire polypeptide encoded by the nucleotide sequences set forth in Figures 1 and 2, substantially identical proteins, or immunogenic fragments of the larger polypeptides. Preferably, immunogenic fragments or protein analogs are employed which are capable of immunizing M. leprae-specific T helper cells without activating suppressor cells, thereby avoiding M. leprae-induced immunosuppression observed in lepromatous patients.

Where immunogenic fragments are employed which have molecular weights less than 1000 daltons, antigenicity can be improved by covalent attachment of the fragments, or haptens, to carrier proteins or synthetic polypeptides to provide conjugate immunogens. Suitable carrier proteins include globulin fractions, the serum albumins of various species, hemocyanin, ovalbumin, lactalbumin, thyroglobulin, and fibrinogen. The number of haptens bound to the carrier protein can vary from 2 to 50, depending upon the conditions of conjugation. Preferably, a given carrier has, on average, at least five peptide haptens covalently attached. Generally, higher antibody titers are obtained using conjugates having higher epitope densities.

Peptides are linked to carrier proteins with a spacer group or crosslinker, of which many suitable examples are known in the art. Among these techniques are the mixed anhydride procedures disclosed by Vaughn et al., J. Am. Chem. Soc. 74:676 (1952), and Karol et al., Proc. Natl. Acad. Sci. USA 57:713 (1967). Alternative conjugation reagents include carbodiimide reagents, for example, 1-ethyl-3(3-dimethylaminopropyl)carbodiimide-HCl, (EDAC), or 1-cyclohexyl-3-[2-morpholinyl-(4)ethyl]carbodiimide metho-p-toluenesulfate.

References disclosing suitable procedures useful in conjugating peptides to proteins by carbodiimide methods include Bauminger et al., Methods in Enzymology 70:151 (1980); Goodfriend et al., Science 143:1344 (1964), and *Jaffe* et al., Immunochem. 7:715 (1970). Carbodiimide coupling reagents are commercially available (Biorad Laboratories, Richmond, California, USA) and product literature discloses applicable techniques for use. Peptides can be coupled to carrier proteins by conversion of haptenic amino groups to p-nitrobenzoylamide by reaction with p-nitrobenzyl chloride. The amide derivative is then reduced to a p-aminobenzoyl derivative which can be coupled to proteins by diazotization. References disclosing such conjugation chemistries include Anderer, Biochim. Biophys. Acta 71:246 (1963) and Deodhar, J. Exp. Med. 111:419, 429 (1960). Alternatively, amino-amino conjugations can be effected using glutaraldehyde as a coupling reagent, for example, as disclosed by Reichlin, Methods in Enzymology 70:159 (1980). Alternatively, disulfide bonds between a cysteine residue in the coupling region of the hapten and a second cysteine residue in the carrier protein can be formed using m-maleimidobenzoyl-N-hydroxy succinimide ester (MBS). This reagent can be reacted with free thiol groups on either peptide or carrier protein, to provide an MBS-acylated intermediate which is then reacted with the other component to provide a conjugate. A suitable technique for use of MBS is disclosed by Kitagawa et al., J. Biochem. 94:1165 (1983). Following conjugation, unreacted materials are separated from the conjugated immunogen by extensive dialysis or gel filtration.

To administer the vaccines, compositions can be prepared which comprise a physiologically acceptable carrier or diluent, for example phosphate-buffered saline. To potentiate the antibody response to immunization, such compositions can optionally comprise an adjuvant such as mineral oil, aluminum hydroxide, or a lymphokine such as interleukin-1 or interleukin-2. Immunization is preferably conducted by intramuscular injection of an initial dose of immunogen folloved several weeks later by a booster. An initial inoculation can consist of from 0.01-1 mg of antigen or conjugate, folloved by a booster of approximately half the initial dose.

Example 1: Cloning and Sequencing of 18/16.6 and 15.2 kDa Antigens

Approximately 10^6 bacteriophage plaques (10^5 recombinant plaques) from a recombinant λgt11 library were screened for expression of an epitope of the M. leprae 18/16.6 kDa protein using a monoclonal antibody, L5, as described below. The recombinant M. leprae library was identical to that described by Young et al., Nature 316:450, 1985.

The monoclonal antibody L5 described by Britton et al. (J. Immunol. 135:4171, 1985) (formerly designated L7.15) recognizes a determinant on the 18/16.6 kDa M. leprae protein antigen and apparently does not cross-react with antigens of other Mycobacterial species. This antibody was prepared (as a gammaglobulin fraction) from ascites fluid at a concentration of 29 mg/ml. It was used at a final concentration of 145 µg/ml in Tris-buffered saline (150 mM Tris-HCl, pH 8.0) containing 0.5% (v/v) Tween 20.

To screen the library, bacteriophage plaques were plated on E. coli Y1090(r^-) (Promega Biotec, Madison, WI, USA), recombinant protein expression was induced by overlaying the plates with nitrocellulose filters impregnated with 10 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and then proteins transferred to the filters were screened for the presence of an L5-binding epitope using the method described by Young et al. (Proc. Natl. Acad. Sci. USA 82:2583, 1985). Biotinylated horse anti-mouse IgG (Vector Laboratories, San Diego, CA) and Vectastain ABC reagent (Vector Laboratories) or Streptavidin-biotin-horseradish peroxidase conjugate (Amersham International, Amersham, UK) were used as developing reagents.

Four recombinant clones were isolated and sequenced as follows. DNA from each clone was prepared and purified from lysates of induced lysogens as described by Maniatis (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York, 1983) and digested with EcoRI. M. leprae insert DNA was isolated from 1.0% low melting agarose and subcloned into suitably-digested M13 mp series vectors (Messing, Meth. Enzymol. 101:20, 1983). M13 recombinant clones were sequenced by the dideoxy chain termination method of Sanger at al. (J. Mol. Biol. 143:161, 1980) on both complementary strands, and compressions were checked using gels containing 25% formamide. Sequence information was assembled and analyzed using a modified version of the computer program of Staden (Nucl. Acids Res.

10:4731, 1982).

Four additional clones were isolated by plaque DNA hybridization, using insert DNA from one of the recombinant λgt11 clones as a probe. One of these clones subsequently was found to express the M. leprae 18/16.6 kDa epitope. Although the clones were all of different lengths, the 5' end of all the sequences were identical to one another.

From the sequence overlap of the clones, it was clear that representation of the gene for the 18 kDa protein within the λgt11 library is non-random, probably reflecting the presence of an EcoRI restriction site within the coding sequence in M. leprae DNA which, perhaps because of its surrounding secondary structure, was not adequately protected by methylation during library construction. Thus, it was apparent that the complete coding sequence for the 18/16.6 kDa protein was unlikely to be retrievable in a single clone (or in overlapping clones) from this particular library.

Accordingly, a cosmid library identical to that described by Clark-Curtiss, et al., (J. Bacteriol. 161:1093, 1985) containing partially PstI-digested M. leprae DNA, was screened by nucleic acid hybridization using a 582 bp insert derived from one of the λgt11 clones as a probe. A 3.2 kb BamHI-PstI fragment was isolated from the library, sequenced by colony hybridization techniques (Hanahan and Meselson, Meth. Enzymol. 100:333, 1983), and shown to contain identical sequences to those from the λgt11 clone inserts. The sequences of the two open reading frames found in this cosmid fragment are shown in Figures 1 and 2, together with their predicted amino acid sequences. The predicted molecular weight of the protein encoded by the first open reading frame is 16,607 Da. A 611 bp DdeI fragment (from position 467 to 1077 in Figure 3) containing the entire 18/16.6 kDa coding sequence was subcloned into the SmaI site of the E. coli expression vector pUC18. This construct, when transformed into a suitable host strain, constitutively expressed a single L5-binding molecule having an apparent molecular weight by SDS-PAGE of approximately 17 kDa.

When the complete nucleotide sequence of the 3.2 kb BamHI-PstI M. leprae fragment was scanned for non-random codon usage using the method disclosed by Fickett (Nucl. Acids Res. 10:5303,

1982), a second open reading frame was identified 560 bases downstream from the 18/16.6 kDa gene. The 137 amino acid sequence deduced from the first methionine residue of this reading frame had a predicted molecular weight of 15,196.

Because some M. leprae-specific helper T cell clones have been found to proliferate in response to λgt11 recombinant material containing the carboxy-terminal 111 amino acids of the 18/16.6 kDa protein, the deduced amino acid sequence was analyzed to determine whether it contains any sequences likely to be good helper T cell immunogens. Applying the algorithm of Margalit et al. (J. Immunol. 138:2213, 1987), five short peptides were discovered which were predicted to be able to form amphipathic helices. The most highly-ranked of these sequences is found in the amino-terminal region of the 18/16.6 kDa protein that is not present in material from the λgt11 clones. The size, amphipathicity score, and location of these sequences are set forth in the following table. In the table, amino acids are numbered in accordance with the sequence shown in Figure 1.

Table 1

Amphipathic Peptides from 18/16.6 kDa Protein

Amino acid range	Peptide Size	Amphipathicity Score
8-23	16-mer	37.2
131-148	18-mer	23.9
97-112	16-mer	14.6
26-40	15-mer	11.9
79-98	20-mer	6.4

CLAIMS

What is claimed is:

1. A recombinant expression vector comprising a DNA sequence encoding a polypeptide which is substantially identical to the 18/16.6 kDa and 15.2 kDa protein antigens of Mycobacterium leprae or an immunologically effective subunit thereof.

2. A recombinant expression vector according to Claim 1, comprising a DNA sequence as shown in Figure 1 or an immunologically effective subunit thereof.

3. A recombinant expression vector according to Claim 2, comprising a DNA sequence encoding a peptide subunit comprising all or an immunologically effective subunit of a sequence selected from the group consisting of amino acids 8-23, 131-148, 97-112, 26-40, or 79-98 as shown in Figure 1.

4. A recombinant expression vector according to Claim 2, comprising a DNA sequence encoding a peptide subunit comprising all or an immunologically effective subunit of amino acids 101-116 as shown in Figure 1.

5. A recombinant expression vector according to Claim 1, comprising a DNA sequence as shown in Figure 2 or an immunologically effective subunit thereof.

6. A substantially homogeneous protein composition consisting essentially of a polypeptide which is substantially identical to the 18/16.6 kDa and 15.2 kDa protein antigens of Mycobacterium leprae or an immunologically effective subunit thereof.

7. A composition according to Claim 6, consisting essentially of a polypeptide which is substantially identical to the 18/16.6 kDa protein antigen of Mycobacterium leprae or an immunologically effective subunit thereof.

8. A composition according to Claim 6, consisting essentially of a polypeptide which is substantially identical to the 15.2 kDa protein antigen of Mycobacterium leprae or an immunologically effective subunit thereof.

9. A peptide immunogen comprising all or an immunologically effective subunit of amino acids 101-116 as shown in Figure 1.

10. A peptide immunogen having an amino acid sequence which is substantially identical to a peptide subunit comprising all or an

immunologically effective subunit of a sequence selected from the group consisting of amino acids 8-23, 131-148, 97-112, 26-40, or 79-98 as shown in Figure 1.

11. A diagnostic immunoassay for the presence of antibody to Mycobacterium leprae, comprising (1) contacting a test sample in which antibody is to be detected with a polypeptide which is substantially identical to the 18/16.6 kDa and 15.2 kDa protein antigens of Mycobacterium leprae or an immunologically effective subunit thereof under conditions enabling formation of polypeptide-antibody complexes, (2) removing unbound components of the test sample, and (3) detecting the presence of polypeptide-antibody complexes using means capable of identifying the presence of bound antibody.

12. An immunoassay test kit comprising as one component thereof a polypeptide which is substantially identical to the 18/16.6 kDa or 15.2 kDa protein antigens of Mycobacterium leprae or an immunologically effective subunit thereof.

13. A vaccine composition comprising a polypeptide which is substantially identical to the 18/16.6 kDa and 15.2 kDa protein antigens of Mycobacterium leprae or an immunologically effective subunit thereof and a suitable diluent or carrier.

Figure 1: Sequence of 18/16.6 kDa Protein

```
ATG CTG ATG CGT ACT GAC CCG TTC CGT GAA CTG GAC CGC TTC GCC   45
Met Leu Met Arg Thr Asp Pro Phe Arg Glu Leu Asp Arg Phe Ala   15

GAG CAA GTG TTA GGT ACG TCT GCC CGC CCA GCA GTA ATG CCC ATG   90
Glu Gln Val Leu Gly Thr Ser Ala Arg Pro Ala Val Met Pro Met   30

GAC GCT TGG CGT GAG GGC GAA GAA TTC GTC GTC GAG TTC GAC CTT  135
Asp Ala Trp Arg Glu Gly Glu Glu Phe Val Val Glu Phe Asp Leu   45

CCT GGC ATC AAA GCC GAT TCA CTG GAC ATT GAC ATC GAA CGC AAC  180
Pro Gly Ile Lys Ala Asp Ser Leu Asp Ile Asp Ile Glu Arg Asn   60

GTA GTC ACC GTG CGG GCC GAG CGC CCA GGC GTC GAC CCC GAT CGG  225
Val Val Thr Val Arg Ala Glu Arg Pro Gly Val Asp Pro Asp Arg   75

GAA ATG CTT GCT GCC GAG CGG CCA CGC GGT GTG TTC AAT CGT CAG  270
Glu Met Leu Ala Ala Glu Arg Pro Arg Gly Val Phe Asn Arg Gln   90

CTG GTT CTC GGC GAA AAC CTC GAC ACC GAA CGG ATC TTG GCT TCC  315
Leu Val Leu Gly Glu Asn Leu Asp Thr Glu Arg Ile Leu Ala Ser  105

TAC CAA GAA GGT GTC CTG AAG TTG TCG ATA CCA GTA GCC GAA AGG  360
Tyr Gln Glu Gly Val Leu Lys Leu Ser Ile Pro Val Ala Glu Arg  120

GCT AAA CCG CGC AAG ATC TCC GTT GAT CGT GGC AAC AAC GGA CAC  405
Ala Lys Pro Arg Lys Ile Ser Val Asp Arg Gly Asn Asn Gly His  135

CAG ACC ATA AAC AAA ACC GCA CAC GAA ATC ATA GAT GCC          444
Gln Thr Ile Asn Lys Thr Ala His Glu Ile Ile Asp Ala          148
```

Figure 2: Sequence of 15.2 Kda Protein

```
ATG TCG TCC CTA ACT ACC GAC CTG ATG CTC ACC CAC AGG CAC CTC   45
Met Ser Ser Leu Thr Thr Asp Leu Met Leu Thr His Arg His Leu   15

AAC GAC CGC GGT CAA GTA GCG GCC ACC ATT GAC GAA ATC CTG AAC   90
Asn Asp Arg Gly Gln Val Ala Ala Thr Ile Asp Glu Ile Leu Asn   30

ACA CAT AAG TTA TTC AGC ACC CGT CAC CGG ATT ATC GAC ACC AGC  135
Thr His Lys Leu Phe Ser Thr Arg His Arg Ile Ile Asp Thr Ser   45

GAA AAT GTG GAC AAC GTC ATC GTA ATT GCT GAT CAG CTG TTC GAC  180
Glu Asn Val Asp Asn Val Ile Val Ile Ala Asp Gln Leu Phe Asp   60

GAC CGA GGA GCA ACG ATC GGA ACG TAT GAT TTC TAC ATC GAT GTT  225
Asp Arg Gly Ala Thr Ile Gly Thr Tyr Asp Phe Tyr Ile Asp Val   75

TCG GCC CTA CCG GAA CAA GTG CAC GAG GGC ATC GTC ATA GCG AGA  270
Ser Ala Leu Pro Glu Gln Cal His Glu Gly Ile Val Ile Ala Arg   90

CTA GCC AAG AGC ACC CAA AAC CGA GCG GGA ATC GAG CAG ACC AAG  315
Leu Ala Lys Ser Thr Gln Asn Arg Ala Gly Ile Glu Gln Thr Lys  105

AGC ATG TTG ATG CTG ATC GAT GGT ATA ACC TAC GAC ACC CCG TTC  360
Ser Met Leu Met Leu Ile Asp Gly Ile Thr Tyr Asp Thr Pro Phe  120

AAC CTG CTC AAA TAC GCT GTC CCA GGA AAC CAA CAT CAA GCT GCG  405
Asn Leu Leu Lys Tyr Ala Val Pro Gly Asn Gln His Gln Ala Ala  135

GCT GGC                                                     411
Ala Gly                                                     137
```

INTERNATIONAL SEARCH REPORT

International Application No PCT/US88/04659

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [3]

According to International Patent Classification (IPC) or to both National Classification and IPC

INT. CL. 4th Ed. C07K 13/00;A61K 37/02;
U.S. CL. 530/350;424/88; 435/320,7

II. FIELDS SEARCHED

Minimum Documentation Searched [4]

Classification System [5]	Classification Symbols
US	530/350;424/88;435/320,7

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [6]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [14]

Category * [15]	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No. [18]
X,P	The Journal of Immunology, Vol. 140,No. 2, Issued 15 January 1988(Ecoth)"Antigenic Proteins of Mycobacterium Leprae Complete Sequence of the Gene for the 18-kDa Protein" Pages 597-601(See summary).	1-13
X	The Journal of Immunology,. Vol. 135, No. 6, Issued December 1985 (Warwick)"Mycobacterium Leprae Antigens Involved in Human Immune Repsonses".pages 4171-4177(See summary).	1-13
X	The Journal of Infectious Diseases, Vol.156, No. 5, Issued November 1987 (William)"Mycobacterial Carbohydrate Antigens for Serological Tescing of Patients with Leprosy". pages 763-769(See Summary).	1-13
X,P	Biological Abstract, Vol. 85, Issued 15 June 1988(Britton)"Antigens of Mycobacterium Leprae identified by immunoprecipitation with sera from leprosy and tuberculosia patients" Abstract no. 122146	1-13

* Special categories of cited documents: [15]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search [2]	Date of Mailing of this International Search Report [2]
21 March 1988	18 APR 1989

International Searching Authority [1]

Signature of Authorized Officer [20]

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, ** with indication, where appropriate, of the relevant passages **	Relevant to Claim No **
X,P	Biological Abstract, Vol. 85, Issued 01 April 1988 (Ehrenberg)"Analysis of the Antigenic profile of Mycobacterium leprae: Cross-reactive and unique specificities of human and rabbit antibodies" Abstract no. 68474).	1-13
X,P	Biological Abstract, Vol. 85, Issued 01 February 1988 (Levis)"Mycobacterial carbohydrate antigens for serological testing of patient with leprosy"(Abstract no. 28156).	1-13
X,P	Chemical Abstract, Vol. 109, Issued 07 November 1988 (Columbus, OH.) "Identification of T-cell-activating recombinant antigens shared among three candidate anti-keorist vaccines, killed M. Leprae M.bovis BCG and Mycobacterium w. ". Abstract no. 168510j	1-13
X,P	Chemical Abstract, Vol. 108, Issued 09 May 1988 (Columbus, OH.) "Antigens of Myco-bacterium leprae identified by immuno-precipitation with sera from leprosy and tuberculosis patients" Abstract no. 165813d	1-13
X,P	Chemccal Abstract, Vol. 109, Issued 05 December 1988 (Columbus, OH.) "The use of a'universal' yeast expression vector to produce an antigenic protein of Myco-bacterium leprae"(Booth) abstract no.206177r	1-13
X,P	Biological Abstract, Vol. 87, Issued November 1988 (Booth)"The use of a universal yeast expression Vector To Produce An Antigenic Protein of Mycobacterium-Leprae". abstract no. 16610	1-13